U.S. Department of Energy
Office of Science

BERKELEY LAB

CTBT & Communication Infrastructure to Support a Secure, Scalable, and Robust SensorNet

Deborah Agarwal

Distributed Systems Department
Lawrence Berkeley National Laboratory

U.S. Department of Energy
Office of Science

The Comprehensive Nuclear Test-Ban Treaty

BERKELEY LAB

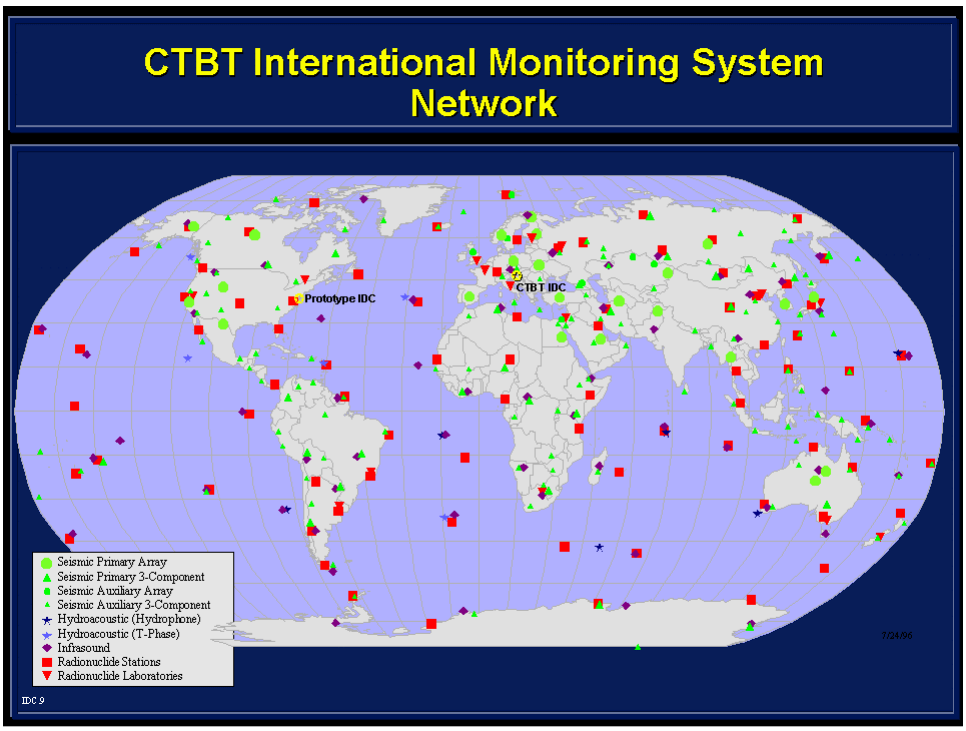
- Bans any nuclear weapon test explosion or any other nuclear explosion
- Drafted at the Conference on Disarmament in Geneva and adopted on September 10, 1996
- Treaty verification requires an ability to detect if a test has been conducted

U.S. Department of Energy
Office of Science

Treaty Verification - Monitoring Compliance

BERKELEY LAB

- International Monitoring System (IMS) 321 stations and 16 radionuclide laboratories
  - Seismic
  - Hydroacoustic
  - Infrasound
  - Radionuclide
- Global Communications Infrastructure
- International Data Center (IDC)
- On-site inspections



U.S. Department of Energy
Office of Science

The International Data Centre (IDC)

BERKELEY LAB

- Collect data from all the sensors
  - Primary/continuous
  - Auxiliary/on demand
- Analyze data
  - automated detection of events
  - reviewed event bulletins
  - determine location of event
- Archive data centrally
- Provide data to all States that request the data

U.S. Department of Energy
Office of Science

The Global Communications Infrastructure (GCI)

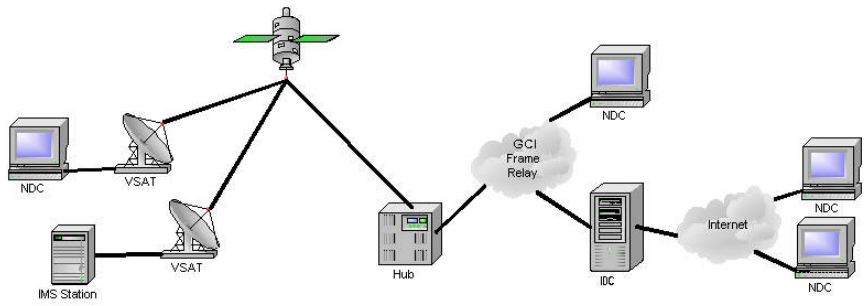
BERKELEY LAB

- Private network
- Built and run by a contractor
- Contract managed by the IDC
- Stations connected by VSAT (Very Small Aperture Terminal earth station)
- 5 satellites with 5 satellite hubs
- Frame relay links connect the hubs to the IDC (land lines)
- Recently added some VPN links over the Internet
- Data distribution using Internet and VSAT
- Global coverage including near the poles

U.S. Department of Energy
Office of Science

The GCI Network

BERKELEY LAB



U.S. Department of Energy
Office of Science

Continuous Data Protocol

BERKELEY LAB

- CD-1.0 implemented by SAIC
  - Push model
    - Sensor connects to the IDC and sends data using TCP
    - If the TCP connection breaks, the sensor starts from the last sent data when a new connection is established
  - Framing
    - Data sent as timestamped frames using compression
    - Data format negotiated at initiation of connection
  - Forwarding
    - Data stored in database and then sent out to any additional receivers

U.S. Department of Energy
Office of Science

Data Reliability and Scalability

BERKELEY LAB

- Data forwarding operations are not reliable
  - With forwarding, if any site in the forwarding chain is down then none of the subsequent sites get the data
  - Hard to reconfigure data paths to move to an alternate primary receiver site
- TCP some problems on high bandwidth delay product links
- No method of requesting corrupted frames
- Difficult to retrieve particular segments of data from storage

U.S. Department of Energy
Office of Science

What is Group Communication?

BERKELEY LAB

- Group communication mechanism
  - Provides one-to-many and many-to-many communication
- Efficient dissemination of messages
  - Network-based duplication (when needed)
  - Targeted retransmissions
  - Bandwidth savings
  - Parallel delivery at multiple locations

U.S. Department of Energy
Office of Science

Group Communication

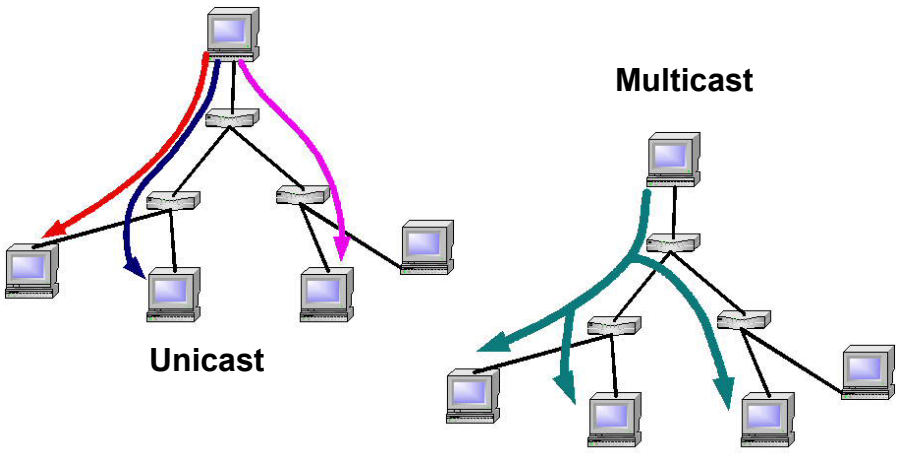
BERKELEY LAB

- Properties similar to TCP
- Application-level program (runs on end systems)
- Uses IP Multicast as the underlying communication mechanism
- Reliable and ordered delivery of messages within a group (negative acknowledgments and retransmissions)
- Tracks group membership

U.S. Department of Energy
Office of Science

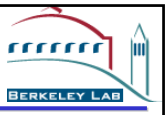
IP Multicast Communication

BERKELEY LAB



U.S. Department of Energy
Office of Science

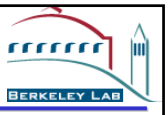
Internet Protocol (IP) Multicast



- Efficient group communication mechanism
  - provides one-to-many communication
  - Best-effort delivery to the group members (unreliable)
- Implemented in the network routers and hosts
  - Class D addresses used for multicast (224.x.x.x - 239.x.x.x)
  - Network components manage routing and duplicate the message as needed
  - Co-exists with TCP and UDP communication mechanisms

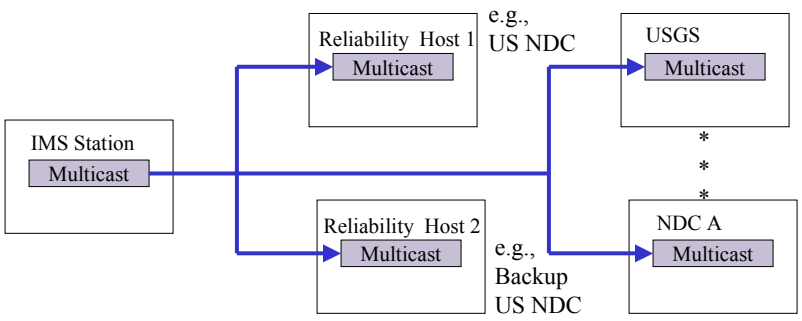
U.S. Department of Energy
Office of Science

CD-1.1 Multicast Capability

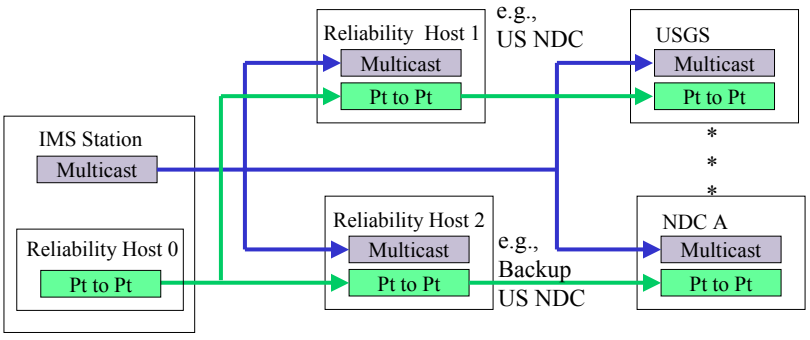


- Motivation
  - efficient transmission of data to multiple sites
  - reliability (not dependent on a single data forwarding site)
  - allow as-needed use of multicast
- Designed and implemented by SAIC
  - data provider multicasts data and provides retransmissions
  - reliability host requests any required catch-up from data provider (unicast)
  - reliability host responsible for catch-up of data consumers (unicast)
  - easy to use either unicast or multicast

CD-1.1 Reliable Multicast - Normal Operation



CD-1.1 Reliable Multicast - Catch-Up Operation



CD-1.1 (latest version)

- Connections
  - Unicast – TCP and UDP
  - Multicast – reliable and unreliable
- Push and pull model
  - Unicast still uses push
  - Multicast uses pull
- Reliability
  - Retransmission requests for missing frames (frame cache)
  - Begin transmission with a short look back
- Format
  - Data format information included in each frame
- Authentication and data integrity
  - Frames are signed at the sensor using PKI
- Authorization
  - Access control lists

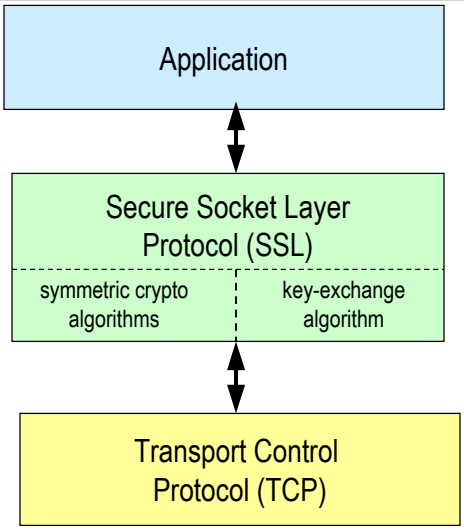
Sensornet vs CTBT

- CTBT
  - data integrity and authenticity is a primary concern
  - data sent unencrypted and signed
  - privacy of analysis products important
  - sensors well-known in advance
  - politics often dominate design decisions
- SensorNet
  - cell towers
  - uplinks serve multiple types of sensors
  - sensors co-located with minimal infrastructure
  - encryption of data

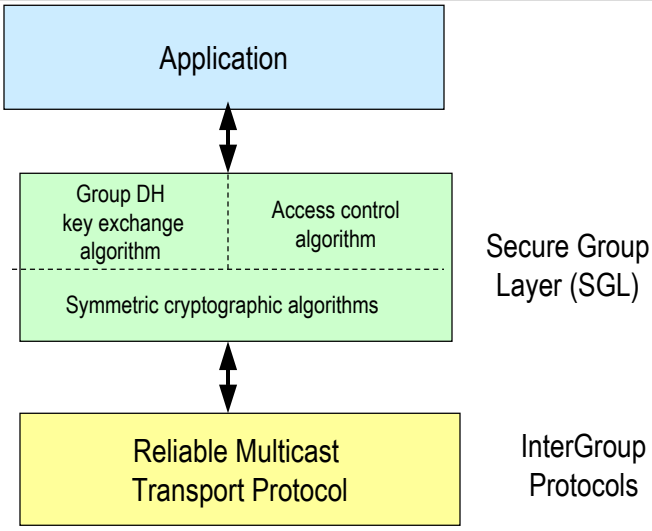
Group Communication Research at LBNL

- Provide an efficient and reliable communication between participants aggregated into groups
  - communication channel directly connecting the participants (no intermediary server)
  - remove dependence on centralized servers (bottleneck, scalability)
  - support participants spread across the Internet
  - support self-organization/self-discovery
  - continue despite network partitions and failures
- Provide secure communication among the participants
  - support confidentiality, authenticity, and integrity
  - support access control based on certificates and passwords

Reliable and Secure Unicast Communication : Architecture



Reliable and Secure Group Communication : Architecture

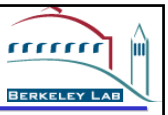


InterGroup Protocol

- All members of the group can send messages to the group
- All processes in the group receive the messages sent within the group
- Membership tracking with notification of membership changes
- Messages delivered to each member of the group in a consistent order
  - total order (timestamp)
  - preserve causality
  - membership changes ordered with respect to messages

U.S. Department of Energy
Office of Science

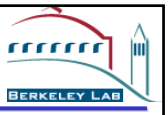
InterGroup Scaling



- Split group into a sender and a receiver group
  - sender group membership
    - processes are in the sender group only while transmitting messages
    - strictly maintained
    - very dynamic (small and fast)
  - receiver group membership
    - semi-anonymous (hierarchical structure)
    - not strictly maintained (RTCP-like)
    - used for retransmissions and garbage collection
    - proxy send for low frequency senders

U.S. Department of Energy
Office of Science

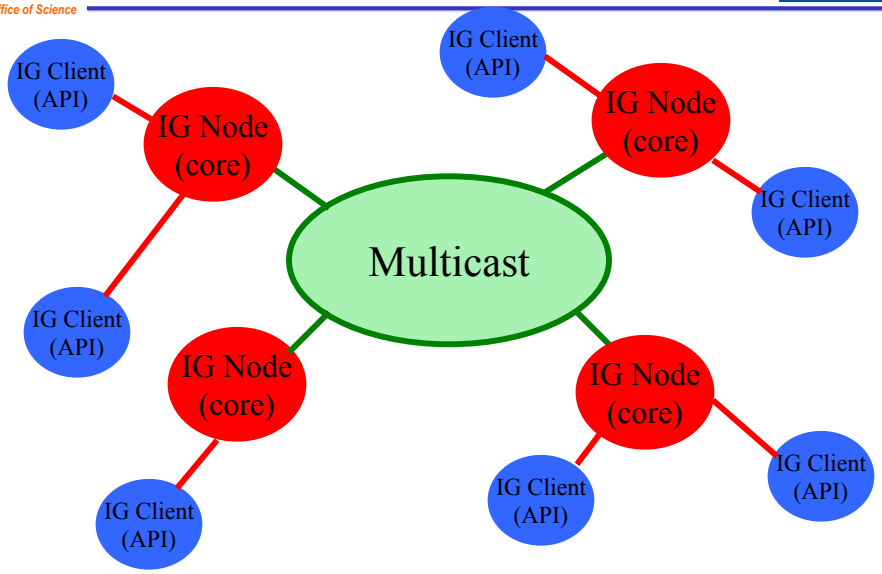
InterGroup Design



- InterGroup node (currently written in Java)
  - core InterGroup capabilities
  - automatically handles membership, message ordering and retransmission of missed messages
  - uses IP Multicast to transmit messages
- InterGroup client API
  - library that connects applications to an InterGroup node
    - locally using unix sockets
    - remotely using TCP
  - TCP connection allows machines without multicast connectivity to participate in InterGroup
  - easy to port to other programming languages

U.S. Department of Energy
Office of Science

InterGroup Architecture



U.S. Department of Energy
Office of Science

Implementation

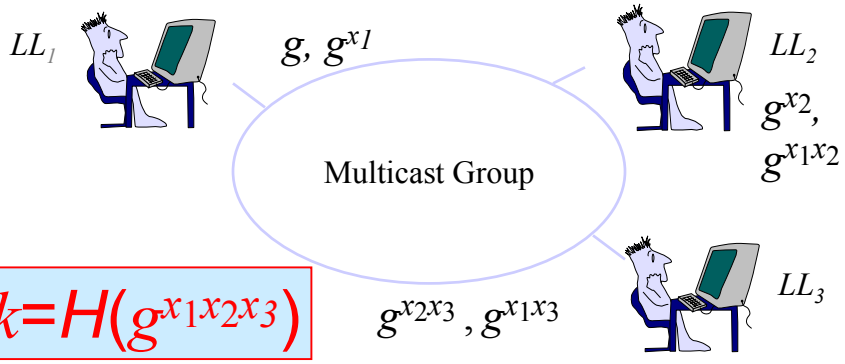
- **Current release v1.5**
  - **IG Node (Java)**
    - deamon listening for client connections
    - all processes in the sender group
    - flow/congestion control very crude
    - reliable group ordered delivery
  - **IG Client (Java, C++, Python)**
    - connects to IG node using TCP
    - C++ client for Unix flavors
      - ⇒ Windows client prototype available
    - Python client is prototype
      - ⇒ SWIG wrapping of the C++ client
- Berket et al., A Practical Approach to the InterGroup Protocols", J. of Future Generation Computer Systems, 2002

The Secure Group Layer: SGL

- A group Diffie-Hellman key exchange algorithm enables group members to establish a session key
- Symmetric crypto algorithms (e.g. DES and HMAC)
  - implements a secure channel
- An access control mechanism makes sure that only the legitimate parties have access to the session key
  - certificate-based
  - password-based
- Provable security

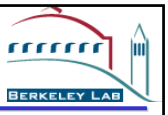
Model of Communication

- A multicast group consisting of a set of n players
  - each player holds a long-lived key (LL)
  - long-lived keys are either a password or a public/private key pair (i.e. PKI)



U.S. Department of Energy
Office of Science

Key Agreement Algorithms

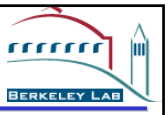


- Support
  - Authentication
    - Password
    - Certificate
    - Anonymous
  - Dynamic membership changes
  - Low power devices
- Security goals
  - Forward secrecy
  - Mutual authentication
  - Secure against dictionary attacks
- Provable security

  Chevassut et al., The Group Diffie-Hellman Problems, Proc of Selected Area in Cryptography, 2002

U.S. Department of Energy
Office of Science

Implementation



- Built in C++
- Framework and interface designed
- Static anonymous and password key exchange prototype implementations complete
- Crypto algorithms implemented for static and dynamic key exchange
- Demonstration chat application implemented

- Agarwal et al., An Integrated Solution for Secure Group Communication in Wide-Area Networks, Proc of IEEE Symposium on Computers and Communication'01

Group Communication in SensorNet

- Sensor data reaches multiple receivers with a single transmission (fault tolerance)
- Allows dynamic ad hoc configuration/addition of new sites
- Scalable to large number of sites/groups/latencies
- Security
  - distributed key agreement
  - password/certificate/anonymous authentication
  - algorithms for low power devices

Incremental Trust Motivation

- Ability to access from anywhere including Internet cafés
- Low threshold for entry into the system
  - Incorporate new users easily
  - Small amount of software downloads
  - No waiting for authorization to enter the system
- Components able to require only the level of authentication and authorization they need. E.g.
  - Weak or no authentication to enter the lobby
  - Strong authentication and authorization for sensitive actions
- Minimize dependence on servers (particularly while the collaboration is small in number)

U.S. Department of Energy
Office of Science

Authentication Model

BERKELEY LAB

- A user has multiple means of authentication
- Authentication for a particular session based on
  - Location
  - Methods available
  - Security of local machine
  - Availability of connection to servers
- Authentication method for a session a property of a user's session
- Authentication method considered in authorization

U.S. Department of Energy
Office of Science

Crossing the borders

BERKELEY LAB

- Escort
  - Chaperone a user in an area they are not normally authorized to access
  - Only provides privileges of the host or less
  - Host able to control the guest's access
- Vouching
  - A user vouches for a less privileged user
  - Temporarily elevates privileges of the vouchee
  - Vouchee able to act without escort

U.S. Department of Energy
Office of Science

Incremental Trust and SensorNet

BERKELEY LAB

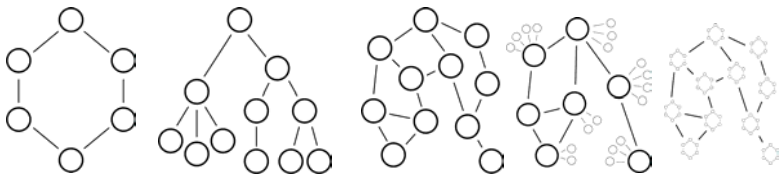
- Allow you to support devices and sites with varying levels of trust
- Developing for collaboratories
- Allow multiple authentication methods
- Quickly provide a means of access for new or temporary users
- Recognize in the system that there are different access authorization levels

U.S. Department of Energy
Office of Science

Peer-to-Peer I/O (P2PIO)

BERKELEY LAB

- Find, access and aggregate information/resources
  - in a large distributed system
  - composed of many autonomous data sources
  - dynamic and heterogeneous information, resources, participants, network



- Example Query
  - "Find sensor data X in the D.C. area and subscribe to updates on changes to the sensor data"

U.S. Department of Energy
Office of Science

P2PIO Features

- Query searches for all items matching a set of criteria
- Provide a means of incrementally searching through a network
  - Iterative access to the result set
  - Can request single or multiple message response to a receive
- Transport independent
- Variety of response mechanisms

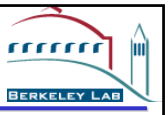
U.S. Department of Energy
Office of Science

General Purpose P2PIO Infrastructure

- Data independent
  - Any structured or semi-structured XML data
  - Integrate static data from relational or XML database
  - Integrate dynamic data generated on-the-fly
- Query language independent
  - Allows for XQuery, XPath, SQL, custom query languages
- Network and Transport independent
  - E.g. TCP, Web Service/SOAP, InterGroup/SGL
  - Arbitrary topologies (allows for hybrid system of systems)
- Fault-tolerant, scalable, interoperable
- Easy-to-use and deploy, easy to extend and customize
- App developers focus on app-specific problems (plug-in)

U.S. Department of Energy
Office of Science

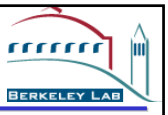
P2PIO Status



- Early version of specification
- Proof-of-concept implementation
  - SOAP-based messaging
  - Accessible as a Grid Service
  - Network support
    - Static implemented
    - Dynamic unicast P2P and InterGroup/SGL in development
  - XQuery and XPath plugins built-in

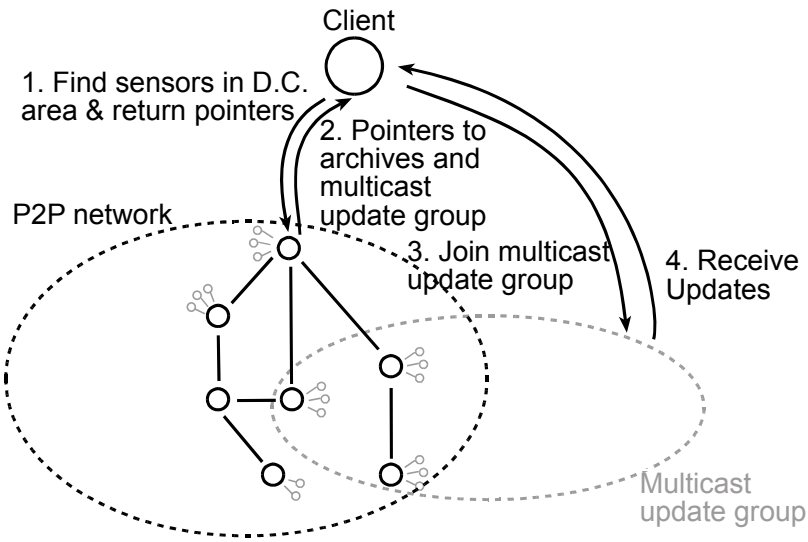
U.S. Department of Energy
Office of Science

SensorNet Query for Updates: P2PIO with InterGroup/SGL



- Client asks for specific sensor information in area of interest and asks for updates
- Local aggregators in charge of number of sensors
- Aggregators are peers in P2PIO and for each query with matching result(s) answer with
  - Pointers to archives of the sensor data
    - Allows access to past data
  - Pointer to SGL group where sensor updates are sent
    - Allows all to receive new data

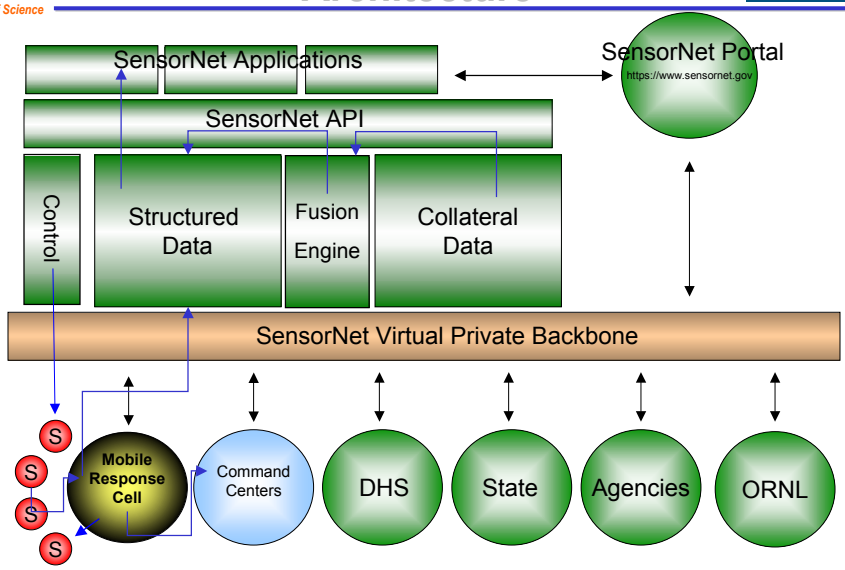
SensorNet Query for Updates:
P2PIO with InterGroup/SGL



SensorNet – Functional Requirements

- **SensorNet Characteristics (from brief)**
  - **A System of Systems Infrastructure**
  - **Real Time Knowledge/Near Real Time Response**
  - **Integration Of Many Dissimilar Sensor Systems**
  - **Scalability To Cover The North American Continent and Hawaii.**
  - **Peer-To-Peer and Conferencing Framework**
    - Near-Simultaneous, Interactive Availability to Data and Services
  - **High Reliability**
    - Self-Organizing/Self-Healing Network Connectivity
    - Distributed Processing
    - Distributed Information Storage
  - **Information Assurance**
    - Encrypted Communication
    - Trust Architecture
      - ⇒ Multi-Level Security
      - ⇒ Access Control
  - **Fusion Of Information Into A Common Operational Data Base and Picture.**

High Level SensorNet Architecture



Further Information

- **LBNL URLs**
  - **http://www-itg.lbl.gov/**
  - **http://www-itg.lbl.gov/CIF/GroupCom/**
  - **http://www-itg.lbl.gov/P2P/file-share/**
- **LBNL Contacts**
  - **DAAgarwal@lbl.gov**
  - **KBerket@lbl.gov**
  - **WHoschek@lbl.gov**
  - **OChevassut@lbl.gov**
- **CTBTO PrepCom WWW site**
  - **http://www.ctbto.org/**